Draft version November 18, 2021
Typeset using LaTeX default style in AASTeX631

Are We Doing Enough: An Analysis of Racial Progress Within Astronomy*

Kevin B. Moposita[1]

[1] *Villanova University*
800 Lancaster Avenue
Villanova, PA 19085, USA

1. ABSTRACT

The purpose of the study was to determine which factor is the most influential towards a student's decision to persist within the astronomy field of study and how that relates to the prevalent racial gap within the department. The 2020 report from the American Institute of Physics contained a survey that compared undergraduate African American and White students through a series of questions. The parameters from that report was utilized for this study and the final calculations and results will be provided once it is available.

2. INTRODUCTION

Among the other fields of science, astronomy ranks as one of the worst departments relative to racial diversity among their demographics. Though there should be a clear reflection of the general population's racial percentage, this is unfortunately not the case. The *2007 Nelson Diversity Survey* highlights the racial disparity among the faculty of the top 50 astronomy schools across the United States. Taking into account every faculty member, 90% identified as White, and approximately 1% identified as Black or Latinx. Analyzing the U.S. Census that year, approximately 66% identified as White, 12.2% as Black, and 15% as Latinx. Instead of reflecting the corresponding percentage, one racial group was much more represented, while the under-represented groups faced a severe lack of representation. The tables of these data sets can be found below. *Figure 1* is representative of the results of the *2007 Nelson Diversity Survey*. *Figure 2* is of the U.S. Census throughout the years along with a racial breakdown of the general population.

Table 4. Tenured/Tenure Track Faculty at the Top 40 Astronomy Departments by Race/Ethnicity, by Gender, and by Rank (FY 2007)*

	White				Black				Hispanic				Asian				Native American				Total
University	Full	Assoc	Asst	*Tot*	Full	Assoc	Asst	*Tot*	Full	Assoc	Asst	*Tot*	Full	Assoc	Asst	*Tot*	Full	Assoc	Asst	*Tot*	
Arizona	16.003	8.001	1	*25.004*	-	-	-	*0*	1	-	-	*1*	-	1	1	*2*	-	-	-	*0*	**28.004**
Johns Hopkins	14.001	-	-	*14.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**14.001**
UC Santa Cruz	17.003	2	1.001	*20.004*	-	-	-	*0*	-	-	1	*1*	3	-	-	*3*	-	-	-	*0*	**24.004**
Chicago	26	6.001	2	*34.001*	-	-	-	*0*	-	-	-	*0*	-	-	1.001	*1.001*	-	-	-	*0*	**35.002**
Cornell	20.001	2	3.001	*25.002*	-	-	-	*0*	-	-	-	*0*	-	1	-	*1*	-	-	-	*0*	**26.002**
Colorado	14.001	4.001	4.001	*22.003*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**22.003**
Hawaii Manoa	24.004	3	2	*29.004*	-	-	-	*0*	-	-	-	*0*	2.001	1	1	*4.001*	-	-	-	*0*	**33.005**
MIT	13.002	3	5	*21.002*	-	-	-	*0*	-	-	1	*1*	-	1	-	*1*	-	-	-	*0*	**23.002**
UT Austin	14.001	1.001	2	*17.002*	-	-	-	*0*	-	-	-	*0*	1.001	-	2.001	*3.002*	-	-	-	*0*	**20.004**
Penn State	10.001	2	2	*14.001*	1.001	-	-	*1.001*	1	-	-	*1*	-	-	-	*0*	-	-	-	*0*	**16.002**
Maryland College Park	9.001	5	1	*15.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**15.001**
UC Berkeley	11.001	1	2	*14.001*	1	-	-	*1*	-	-	-	*0*	1.001	1	-	*2.001*	-	-	-	*0*	**17.002**
Massachusetts Amherst	9.002	4.002	2	*15.004*	-	-	-	*0*	-	-	-	*0*	-	3	1	*4*	-	-	-	*0*	**19.004**
CA Institute of Tech.	11.003	-	2.001	*13.004*	-	-	-	*0*	-	-	-	*0*	1	1	-	*2*	-	-	-	*0*	**15.004**
Wisconsin	5.002	3.001	3.001	*11.004*	1	-	-	*1*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**12.004**
Columbia New York	9.002	4.002	7.002	*20.006*	-	-	-	*0*	-	-	-	*0*	-	2.001	-	*2.001*	-	-	-	*0*	**22.007**
UC San Diego	9.002	-	3	*12.002*	-	-	-	*0*	-	-	-	*0*	1	-	-	*1*	-	-	-	*0*	**13.002**
Princeton	14.002	-	2.001	*16.003*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**16.003**
Illinois Urbana-Champaign	5	4	3	*12*	-	-	-	*0*	-	-	-	*0*	1.001	-	1	*2.001*	-	-	-	*0*	**14.001**
Ohio St	9.001	3.001	3.001	*15.003*	-	-	-	*0*	-	-	-	*0*	1	1.001	-	*2.001*	-	-	-	*0*	**17.004**
Harvard	14.001	1	2	*17.001*	-	-	-	*0*	1	-	-	*1*	1	-	1.001	*2.001*	-	-	-	*0*	**20.002**
Washington	8.002	1.001	2	*11.003*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**11.003**
Florida	9.001	4.001	1	*14.002*	-	-	-	*0*	-	-	1	*1*	1	-	-	*1*	-	-	-	*0*	**16.002**
SUNY Stony Brook	7	-	1	*8*	-	1	-	*1*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**9**
Minnesota	8.001	-	-	*8.001*	1	-	-	*1*	-	-	-	*0*	-	1.001	-	*1.001*	-	-	-	*0*	**10.002**
Virginia	9	1	2.001	*12.001*	-	-	-	*0*	-	-	-	*0*	1	1	-	*2*	-	-	-	*0*	**14.001**
Michigan	5	-	8.004	*13.004*	-	-	-	*0*	1.001	-	-	*1.001*	-	-	-	*0*	-	-	-	*0*	**14.005**
Pittsburgh	4.001	2	-	*6.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**6.001**
Rochester	7.001	1.001	-	*8.002*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**8.002**
Iowa	2	2	1.001	*5.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**5.001**
New Mexico St	4	2	2.001	*8.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**8.001**
Yale	7	1	-	*8*	-	-	-	*0*	-	-	-	*0*	1.001	1.001	-	*2.002*	-	-	-	*0*	**10.002**
Indiana	5.002	1	1.001	*7.003*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**7.003**
Boston	11	3.001	1.001	*15.002*	-	-	-	*0*	-	-	-	*0*	1	-	-	*1*	-	-	-	*0*	**16.002**
Arizona St	4	1	1	*6*	-	-	-	*0*	-	-	-	*0*	-	1.001	-	*1.001*	-	-	-	*0*	**7.001**
Rice	1	2	4	*7*	-	-	-	*0*	-	-	-	*0*	1	-	-	*1*	-	-	-	*0*	**8**
Southern California	2	1.001	-	*3.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**3.001**
Case Western Reserve	3.001	-	1.001	*4.002*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**4.002**
Delaware	9	1	2	*12*	-	1.001	-	*1.001*	-	-	-	*0*	1	-	-	*1*	-	-	-	*0*	**14.001**
MS State	2	-	1.001	*3.001*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	-	-	-	*0*	**3.001**
Astronomy Total	380.043	79.015	80.020	*539.078*	4.001	2.001	0	*6.002*	4.001	0	3	*7.001*	18.005	16.005	8.003	*42.013*	0	0	0	*0*	**594.094**
Percent within race	70%	15%	15%	*100%*	67%	33%	0%	*100%*	57%	0%	43%	*100%*	43%	38%	19%	*100%*	0%	0%	0%	*0%*	
Percent of grand total	64.0%	13.3%	13.5%	*90.7%*	0.7%	0.3%	0%	*1.0%*	0.7%	0%	0.5%	*1.2%*	3.0%	2.7%	1.3%	*7.1%*	0%	0%	0%	*0%*	**100%**
Females in column	11.3%	19.0%	25.0%	14.5%	25.0%	50.0%	0%	33.3%	25.0%	0%	0%	14.3%	27.8%	31.2%	37.5%	31.0%	0%	0%	0%	0%	15.8%

*By astronomy research expenditures FY2004, NSF, www.nsf.gov/statistics/nsf06323/tables.htm#rd7; numbers after decimals designate females.
Reference: "The Nelson Diversity Surveys" Nelson, D. J.: Norman, OK, 2007; http://cheminfo.chem.ou.edu/faculty/djn/diversity/top50.html

Figure 1. The figure is the concluding statistics from the *2007 Nelson Diversity Survey*.

* Released on March, 1st, 2021

Year	Total	White	Black	Hispanic	Asian
				Number (in thousands)	
1980	226,546	180,906	26,142	14,609	3,563
1985	237,924	184,945	27,738	18,368	5,315
1990	248,791	188,315	29,304	22,379	6,996
1995	262,803	193,328	31,590	27,107	8,846
2000	282,158	195,771	34,414	35,629	10,436
2001	284,915	196,325	34,793	36,958	10,777
2002	287,501	196,773	35,147	38,264	11,103
2003	289,986	197,152	35,457	39,579	11,432
2004	292,806	197,727	35,811	40,956	11,782
2005	295,583	198,244	36,145	42,354	12,145
2006	298,442	198,781	36,499	43,777	12,520
2007	301,280	199,272	36,849	45,219	12,901

Figure 2. This figure is of the *2007 United States Census.*.

Similarly, the demographics of graduate level students reflect this disparity. From 2002 to 2012, under-represented minorities only made up approximately 3% of total PhDs awarded for astronomy Rudolph et al. (2020). In the span of a decade, the percentage of PhDs conferred to under-represented minorities did not surpass even 4% of the entire pool of graduate students. If the disparity persists as early as the graduate school level, how might the undergraduate level pool look like? Assuming the undergraduate pool is more diverse than both the graduate level and full-time faculty, this may be interpreted as the 'turn-off point' in which these under-represented groups deviate from pursuing a more advanced position in this field. If this is true, it would be indicative of issues occurring between undergraduate and graduate school. What is going on during this period?

Just as analysis of the racial demographics of full-time faculty and graduate students were conducted, the diversity breakdown of undergraduate students within astronomy will also be observed. As suspected, the racial disparity is not as severe in the undergraduate level. There is a higher percentage of Black and Latinx people earning bachelors degrees than in the two previous areas. The trend within both of these graphs indicate that the percentage will only increase. One important note is that the percentage of Black bachelor recipients very slowly increases American Physical Society (2020a) while the Latinx percentage has increased at a much faster rate American Physical Society (2020b). *Figure 3* displays the percentage of Black bachelor recipients while *Figure 4* does the same for the Latinx group.

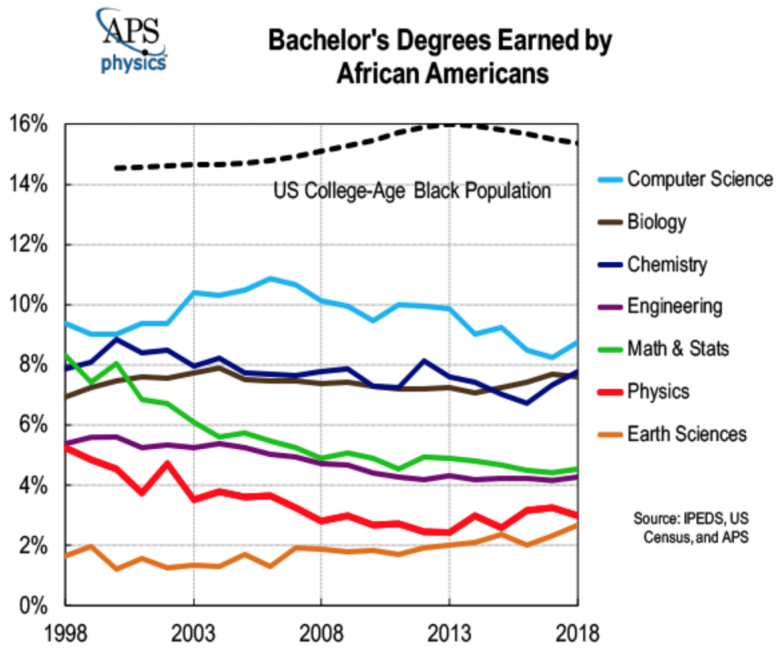

Figure 3. The graph is from the report of the *American Physical Society.*

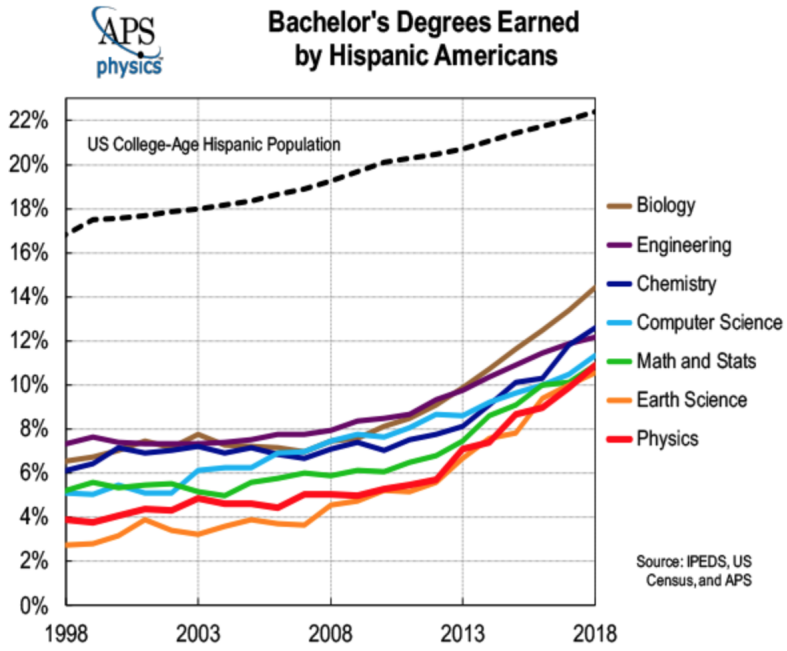

Figure 4. This graph is also from the report of the *American Physical Society*.

With this information in hand, institutions across the country have been implementing social programs within their respective astronomy departments and creating task forces meant to place importance on diversity and inclusion. However, how effective have they been? Just as it is important to implement these programs, it is also important to keep up with them and analyze their effectivity.

For this project, two objectives are being set: determining where the large racial gap occurs along an astronomer's path to tenure and the effectivity of programs set in place by institutions that are meant to highlight diversity and inclusion. There is the potential of a connection between the two objectives. Depending on the effectivity of these programs, it may either deplete or contribute towards the present day racial gap.

3. METHODOLOGY

3.1. *Background*

For this study, factors that potentially contribute to the racial gap within the undergraduate level needed to be determined. Along with their determination, it was of vital importance that each one be introduced with equal importance. Failure to do so would not only introduce a level of bias, but skewer the accuracy of the results. The question quickly became, *how could this issue be prevented?* Past this issue, the idea behind defining said factors was to determine if correlation exists between each one and the mentioned undergraduate diversity gap. Assuming correlation were to exist, it would provide an partial answer as to why the gap exists.

3.2. *Data Retrieval*

The data set from the AIP TEAM-UP Team (2020) of the *American Institute of Physics* was utilized. Alongside the data set, the report also provides a list of factors that are believed to be potential causes to the under-representation of Black undergraduate students within physics and astronomy. Not only is the inclusion of each factor thoroughly explained, but is also backed up by various literature. In light of this, the factors the report has listed will also be the factors utilized in this project. It is also recognized that these factors are not meant to be representative of all possible ones, but of the ones with significant background. Due to the survey containing significant data only for White and Black students, the scope of this project will shift to these two racial groups. *Figure 5* displays the conducted survey from this report. The data set itself was collected through means of a survey in which 187 undergraduate students answered questions regarding their college experience. A Likert scale was utilized for the participants to record their answers. The AIP report presented each racial group's mean response according to the question and factor it was accounting for. However, to check for correlation of each factor, a linear regression model must be in place. Statistical

tests are then performed with the responses to determine the existence of correlation. Unfortunately, the individual response would be required to carry these tests out and not the group's average response. While the report mentioned a the original responses for a couple of the questions, it did not do so for all of them. To solve this issue, an email was sent to the team responsible for AIP's report in an effort to obtain the original, unedited student responses.

Barring a response from AIP, a temporary solution was implemented to continue carrying out this study. Though AIP's report unfortunately does not provide individual responses from their student survey, they do provide the mean average and standard deviation of each survey question from each ethnic group that was involved in the study. With this information, python was utilized to create individual mock student responses that followed those constraints. With those limits imposed, it allows for the most realistic individual responses available. To maintain consistency with the actual survey, the total number of responses was also kept the same. *Table 1* contains the parameters that were utilized in generating the mock data points.

Factor/indicator	Overall N=167		African Am (AA) n= 52		Black-Biracial (BB) n = 32		White (W) n = 53		Other (O) n = 30		F-test	p-value	Significant Mean Comparisons (Bonferroni)
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.			
Intent to persist													
Completing major at institution	3.66	0.64	3.69	0.54	3.72	0.63	3.70	0.61	3.47	0.82	1.10	0.35	None
Intent to withdraw													
Transferring to another inst.	1.88	0.96	2.04	0.95	1.71	0.97	1.73	0.84	2.07	1.10	1.60	0.19	None
Changing majors	1.92	0.91	1.96	0.90	1.81	0.91	1.90	0.89	1.96	0.99	0.21	0.89	None
Leave field after degree	1.79	0.86	1.73	0.75	1.74	0.82	1,82	0.93	1.90	0.99	0.31	0.82	None
Certainty of Major													
Majoring in Physics right choice	3.30	0.78	3.24	0.79	3.09	0.94	3.40	0.71	3.45	0.68	1.12	0.34	None
Important to be a physicist	3.16	0.87	3.24	0.85	2.91	0.94	3.10	0.90	3.40	0.68	1.31	0.28	None
Financial concerns													
Paying for college	2.82	1.10	2.83	1.20	2.78	1.0	2.70	1.07	3.10	1.06	0.95	0.42	None
Working interf. studies	2.46	1.13	2.44	1.18	2.41	1.2	2.47	1.03	2.50	1.97	0.04	0.99	None
Paying college debt	2.60	1.21	2.79	1.18	2.56	1.2	2.64	1.18	2.90	1.24	2.46	0.06	None
Faculty interactions													
Phys fac encourage class part	4.07	1.13	4.08	1.20	4.25	0.84	3.63	1.40	3.63	1.40	2.07	0.11	None
Phys fac interested in my ideas	4.02	1.20	4.06	1.18	4.25	0.98	4.09	1.15	3.59	1.48	1.76	0.16	None
Comfortable approaching fac	4.16	1.09	3.94	1.19	4.35	0.84	4.36	0.90	3.97	1.38	1.94	0.13	None
Fac affirm ability to do physics	3.10	0.90	3.08	0.86	3.22	0.79	3.13	0.94	2.93	0.99	0.56	0.64	None
Classroom Self-Efficacy													
Confident on physics assignments	4.06	0.98	3.90	1.14	4.03	0.78	4.28	0.89	3.97	0.99	1.47	0.23	None
Doing excel job physic exams	3.66	1.14	3.42	1.25	3.69	0.93	3.94	1.05	3.57	1.25	1.93	0.21	None
Confident handling lab equip	4.05	1.10	4.08	1.10	4.06	1.17	4.00	1.36	4.07	1.01	0.03	0.99	None
Self-efficacy as a physicist													
See oneself as physicist	3.36	0.74	3.37	0.71	3.22	0.71	3.43	0.82	3.37	0.67	0.57	0.64	None
Others regard one as physicist	2.96	0.90	3.00	0.95	2.97	0.82	3.02	0.91	2.80	0.92	0.42	0.74	None
Awareness of career opport.	2.83	0.88	2.98	0.89	2.72	0.89	2.85	0.91	2.63	0.76	1.20	0.31	None
Learning strategies													
Seek help from peers	3.26	0.87	3.15	0.94	3.41	0.61	3.28	0.86	3.27	1.01	0.56	0.64	None
Seek help from professor	3.04	0.87	3.10	0.87	3.00	0.92	3.08	0.85	2.90	0.88	0.38	0.77	None
Seek help from online resources	3.50	0.70	3.59	0.69	3.50	0.57	3.43	0.74	3.47	0.78	0.50	0.68	None
Departmental belonging													
Belonging academic dept comm	3.97	1.22	3.96	1.10	3.90	1.28	4.13	1.24	3.76	1.35	0.61	0.61	None
Community with peers in major	3.86	1.26	3.69	1.24	3.69	1.25	4.15	1.12	3.67	1.47	1.51	0.21	None
Departmental supportive env.	4.15	1.03	4.12	1.09	4.38	0.87	4.30	0.85	3.73	1.28	2.59	0.05	O < W**; O < BB**
Sense of community with peers of same ethnic group	3.53	1.31	4.04	1.07	3.21	1.21	3.66	1.33	2.77	1.33	7.63	0.01	BB< AA**; O< W**; O<AA**
Pro-social behaviors													
Organizations that improve soc	3.39	0.67	3.50	0.70	3.09	0.77	3.42	0.57	3.50	0.57	2.97	0.03	BB < AA**; BB < O*
Making the world a better place	3.49	0.68	3.58	0.64	3.22	0.87	3.62	0.56	3.40	0.62	2.97	0.03	BB < W**; BB<AA*
Benefit own community	3.29	0.88	3.65	0.59	3.16	0.92	3.21	0.95	3.00	0.98	4.73	0.03	W<AA**; BB<AA*;O<AA***
Mentor others in the major	3.51	0.79	3.62	0.69	3.28	0.81	3.55	0.64	3.50	0.68	1.59	0.19	None
Perceptions of prejudice													
Treated negative in class & labs	1.55	1.09	1.92	1.33	1.21	0.49	1.33	0.97	1.57	1.07	3.79	0.01	W<AA**; BB<AA**
Seen other treated negatively	1.57	1.03	1.78	1.30	1.23	0.50	1.52	0.84	1.23	0.50	1.88	0.13	None
Academic Performance													
Overall	3.50	0.74	3.31	0.73	3.32	0.65	3.81	0.74	3.47	0.68	5.35	0.01	AA<W**; BB<W**
In the major	3.51	0.82	3.30	0.81	3.26	0.86	3.85	0.75	3.57	0.73	5.45	0.01	AA<W**; BB<W**

Notes: * p < .10; ** p < .05; *** p < .01

Figure 5. The results of the survey conducted by AIP

3.3. *Data Processing*

As briefly mentioned, the obtained data would be utilized to run linear regression models to determine correlation for each factor. Before the next steps could be taken, it was important to make sure the survey was impartial. To ensure this, the 7 Classical Assumptions of Ordinary Least Squares was consulted. The purpose of these assumptions is meant to act as a checklist that the survey data must satisfy so as to yield the best possible estimates. Following this approval, the BREUSCH (1978), Breusch & Pagan (1979), and Ramsey (1969) are the tests that will aid in completing the objective. The Breusch-Pagan test will make sure the error within the data is normally distributed. The Breusch-Godfrey test will establish if correlation is present with each factor. The RESET test is meant to act

as a diagnostic for correctness of a functional form. With each factor undergoing all three tests, all aspects will be covered when determining if correlation is present. All of the statistical tests were carried out in Python.

As results begin to come in, I will continue to add more onto this section.

4. DISCUSSION

With the mock survey results, the mentioned linear regression tests were performed to statistically determine which factor is most influential when undergraduate students pondered what would cause them to drop their pursuit of a bachelor's degree in astronomy.

4.1. *Possible Solution*

As previously mentioned, statistical analysis indicate that *FACTOR* is the most influential when undergraduate students think about potentially dropping astronomy as a major. In light of this assertion, the question then becomes *what can be done to combat this issue?*

A possible solution that institutions across the country may want to begin to think about will be mentioned once final results become available.

5. CONCLUSION

In conclusion, as a result of the tests performed on the survey results, *FACTOR* is the one that has the most statistical influence when undergraduate students ponder whether they will continue their pursuit of a bachelor's degree in astronomy. In response to this, institutions may want to think about implementing *ACTION*. As supported by *LITERATURE*, an implementation of this kind of resource would have a profound effect on the students who are being the most impacted. Though it will not completely close the present nationwide racial gap within the field of study, it will provide a sense of relief which is meant to begin bridging together the great disparity.

Linear Regression along with the statistical tests are currently being carried out and not yet available.

6. ACKNOWLEDGMENT

Completion of this project would not be possible without the guidance of Dr. Andrej Prsa and Dr. Scott Engle. Both professors were very instrumental in overcoming obstacles throughtout the duration of the study.

REFERENCES

AIP TEAM-UP Team. 2020, The Time is Now: Systemic Changes to Increase African Americans with Bachelor's Degrees in Physics and Astronomy

American Physical Society. 2020a, Bachelor's Degrees Earned by African Americans, by Major

—. 2020b, Bachelor's Degree Earned by Hispanic Americans, by Major

BREUSCH, T. S. 1978, Australian Economic Papers, 17, 334, doi: https://doi.org/10.1111/j.1467-8454.1978.tb00635.x

Breusch, T. S., & Pagan, A. R. 1979, Econometrica, 47, 1287. http://www.jstor.org/stable/1911963

Ramsey, J. B. 1969, Journal of the Royal Statistical Society. Series B (Methodological), 31, 350. http://www.jstor.org/stable/2984219

Rudolph, A., Basri, G., Agüeros, M., et al. 2020, Bulletin of the AAS, 51. https://baas.aas.org/pub/2019i0101